Studii şi Cercetări Ştiinţifice Universitatea Bacău Seria Biologie
ISSN 1224-919X
Vol. 12, Issue 1, 2007

STRUCTURE VERSUS BIOLOGICAL ROLE OF SUBSTITUTED THIADIAZOLE- AND THIADIAZOLINE- DISULFONAMIDES

Sorana Daniela BOLBOACĂ, Lorentz JÄNTSCHI

ABSTRACT

The relationship between structure and inhibition activity on carbonic anhydrase I of a set of forty substituted 1,3,4-thiadiazole- and 1,3,4-thiadiazoline-disulfonamides has been investigated by using the Molecular Descriptors Family method. The molecular descriptors family has been generated starting with the information obtained from the compounds structure and the descriptors were calculated. The MDF SAR equations were obtained using the molecular descriptors set. Significant models with best performances in estimation were identified using squared correlation coefficient, F-parameter and its significance. The prediction abilities of two multivariate models were analyzed, and the correlation coefficients were compared with the correlation coefficients obtained by previous reported models. The results revealed that the MDF SAR is a useful approach in characterization of inhibition activity on carbonic anhydrase I of studied substituted 1,3,4-thiadiazole- and 1,3,4-thiadiazoline-disulfonamides.

KEYWORDS**:** Molecular Descriptors Family on Structure-Activity Relationships (MDF SAR); Substituted 1,3,4-Thiadiazole- and 1,3,4-Thiadiazoline-Disulfonamides; Carbonic Anhydrase I (CA I); Inhibition Activity

INTRODUCTION

Carbonic anhydrases are ubiquitous metallo-enzymes that catalyze the inter-conversion of the carbon dioxide and the bicarbonate ion, this reaction being fundamental to many processes such as respiration, renal tubular acidification and bone resorption [7]. In human there are known as active types eleven CA isozymes [12]. Carbonic anhydrase I is localized at the level of cytosol, and it is known to have low catalytic activity comparing with carbonic anhydrase II [21] and medium affinity for sulfonamides [6]. Since introduction of the quantitative structure-activity relationships method [10], many researchers investigated by using different descriptors the relationships of the inhibitory activity on CA I of aromatic/heterocyclic sulfonamides [9, 4, 19] and their activities.

A number of forty substituted 1,3,4-thiadiazole- and 1,3,4-thiadiazoline-disulfonamides were previously studied as inhibitors on carbonic anhydrase I [20]. The equations of the best performing QSAR models previously reported are presented in Table 1. The descriptors used were: the polarizability tensor (Π_{xx}, Π_{yy}, Π_{zz}), the dipole moment (μ_x, μ_z), the solvation energy (ΔH_S), the charges on azot atom (Q_{Nr2}), the charges of the atoms of the primary sulfonamide group (Q_{S1}, Q_{O1}), the charges of the atoms of the secondary sulfonamide group (Q_{S2}, Q_{O2}), the charges on specific C atom (Q_{Cr2}), the charges on specific N atoms (Q_{Nr2}), and partition coefficient (LogP).

Note that the model no. 3 was obtained on thiadiazoles and the model no. 4 was obtained on the thiadiazolines compounds.

Table 1. The previous reported models

Model	Expression
1	$\log IC_{50} = 9.29\cdot10^{-3}\cdot\Pi_{xx}-5.72\cdot10^{-3}\cdot\Pi_{zz}-13.04\cdot Q_{Nr2}+17.07\cdot Q_{S1}+1.560\cdot Q_{S2}+6.90\cdot10^{-2}\cdot\mu_x-50.83$
2	$\log IC_{50} = -3.68\cdot10^{-3}\cdot\Pi_{zz}+3.152\cdot Q_{Cr2}+0.157\cdot\mu_x+0.400\cdot LogP-24.62\cdot Q_{O1}-44.1$
3	$\log IC_{50} = 59.43\cdot Q_{S1}+0.1359\cdot\mu_x-0.0300\cdot\mu_z-0.0204\cdot\Delta H_S+98.87\cdot Q_{O1}+27.83$
4	$\log IC_{50} = 8.47\cdot10^{-3}\cdot\Pi_{yy}-5.871\cdot Q_{S2}-1.787\cdot E_H-1.575\cdot E_L+0.0501\cdot\Delta H_S-82.31\cdot Q_{O1}-16.36\cdot Q_{O2}-182.6$

Source: Supuran & Clare, 1999

The statistical characteristics expressed as squared correlation coefficients (R^2), leave-one-out scores (Q^2), standard errors of estimate, Fisher variance ratio (F), and the sample size (n) of models presented in Table 1 were summarized in Table 2.

Table 2. Statistics of the previous reported QSAR models

Model	R^2	Q^2	s	F	n
1	0.753	0.628	0.289	16.78	40
2	0.700	0.570	0.201	13.98	36
3	0.909	0.502	0.18	27.94	20^{*}
4	0.917	0.712	0.21	18.92	20^{**}

* = the thiadiazoles **= the thiadiazolines
Source: Supuran & Clare, 1999

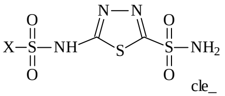
Starting from the hypothesis that there is a relationship between the structure of biological active compounds and their structure, an original method called molecular descriptors family on the structure-activity relationships (MDF SAR) has been developed [13]. The MDF SAR method proved its usefulness in estimation and prediction of inhibition activity on CA IV [14] and CA II [18], and on other activities and properties of active biological compounds [15]. The aim of the research was to study the estimation and prediction abilities of the MDF SAR methodology in modeling of the inhibition activity on carbonic anhydrase I of a sample of forty substituted 1,3,4-thiadiazole- and 1,3,4-thiadiazoline-disulfonamides.

MATERIAL AND METHODS

SUBSTITUTED 1,3,4-THIADIAZOLE- AND 1,3,4-THIADIAZOLINE-DISULFONAMIDES

A sample of twenty 1,3,4-thiadiazole disulfonamides (cle_) and twenty 1,3,4-thiadiazoline (cli_) disulfonamides, with inhibition activity on carbonic anhydrase I was included into the study. The measured inhibitory activity of compounds, expressed in logarithmical scale (as logarithm of concentration of the agent that is required for fifty percent inhibition in vitro - log IC_{50}), was took from a previously reported study [20]. The experimental values expressed in nM, the compounds generic structure, abbreviation and substituent are presented in Table 3.

Table 3. Generic structure of compounds, substituent and associated measured activity


cle_

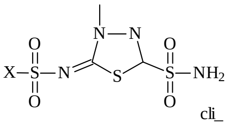

cli_

Abb.	X	log IC_{50}	Abb.	X	log IC_{50}
cle_01	Me	1.0000	cli_01	Me	1.2304
cle_02	PhCH_2	0.8451	cli_02	PhCH_2	0.7782
cle_03	4-Me-C_6H_4	0.6990	cli_03	4-Me-C_6H_4	0.6990
cle_04	4-F-C_6H_4	0.6021	cli_04	4-F-C_6H_4	0.9031
cle_05	4-Cl-C_6H_4	0.6021	cli_05	4-Cl-C_6H_4	0.9031
cle_06	4-Br-C_6H_4	0.4771	cli_06	4-Br-C_6H_4	0.6990
cle_07	4-MeO-C_6H_4	0.6990	cli_07	4-Me0-C_6H_4	0.7782
cle_08	4-AcNH-C_6H_4	1.0000	cli_08	4-AcNH-C_6H_4	0.3010
cle_09	4-H_2N-C_6H_4	0.7782	cli_09	4-H_2N-C_6H	0.0000
cle_10	3-H_2N-C_6H_4	0.9542	cli_10	3-H_2N-C_6H_4	0.0000
cle_11	4-O_2N-C_6H_4	0.4771	cli_11	4-O_2N-C_6H_4	0.9031
cle_12	3-O_2N-C_6H_4	0.3010	cli_12	3-O_2N-C_6H_4	0.8451
cle_13	2-O_2N-C_6H_4	0.6990	cli_13	2-O_2N-C_6H_4	0.6990
cle_14	Me_2N	1.2788	cli_14	Me_2N	0.9542
cle_15	2-$HO_2CC_6H_4$	0.0000	cli_15	2-$HO_2CC_6H_4$	0.0000
cle_16	4-(2,4,6-Me_3Py^+)C_6H_4	1.2553	cli_16	4-(2,4,6-Me_3Py^+)C_6H_4	1.2304
cle_17	4-(2,4,6-Ph_3Py^+)C_6H_4	2.5563	cli_17	4-(2,4,6-Ph_3Py^+)C_6H_4	2.6580
cle_18	2,4-$(O_2N)_2C_6H_3$	1.0792	cli_18	2,4-$(0_2N)_2C_6H_3$	1.0000
cle_19	4-Cl-3-O_2N-C_6H_3	0.9542	cli_19	4-Cl-3-O_2N-C_6H_3	0.8451
cle_20	2,4,6-$Me_3C_6H_4$	1.1761	cli_20	2,4,6-$Me_3C_6H_4$	1.1139

[log IC_{50}] = nM; X = substituent; Me = methyl; Ph = phenyl; Ac = acetyl; Py+ = pyridine

MOLECULAR DESCRIPTORS FAMILY ON STRUCTURE-ACTIVITY RELATIONSHIPS (MDF SAR)

The MDF-SAR method integrate the complex information obtained from the structure of the substituted 1,3,4-thiadiazole- and 1,3,4-thiadiazoline-disulfonamides into models in order to explain the inhibition activity of these compounds on carbonic anhydrase I (CA I). A number of six steps were used into the modeling process [13].

The compounds preparation for modeling was done in the first step. In this step, the three-dimensional structure of substituted 1,3,4-thiadiazole- and 1,3,4-thiadiazoline-disulfonamides were built up by using HyperChem software (http://hyper.com/products/) and the file with measured inhibition on CA I was created.

In the second step, the Molecular Descriptors Family (MDF) was generated and the value of each descriptor was calculated for the studied compounds. The resulted descriptors had a name of seven-letters that explained the modality of its construction [13]: the compound characteristic relative to its geometry (*g*) or topology (*t*) - the 7th letter; the atomic property (which can be: cardinality - *C*, number of directly bonded hydrogen's - *H*, atomic relative mass - *M*, atomic electronegativity - *E*, group electronegativity - *G*, or partial charge - *Q*) - the 6th letter; the atomic interaction descriptor - the 5th letter; the overlapping interaction model - the 4th letter; the fragmentation criterion used in calculations - the 3rd letter [8, 16]; the cumulative method of fragmentation - the 2nd letter, and the linearization procedure applied in generation of molecular descriptors - the 1st letter.

The best performing MDF SAR models were selected in the third step. Three criterion were used: (1) the goodness-to-fit of the model (the correlation coefficient and the squared correlation coefficient; the values closest to ±1 indicated a good model); (2) the co-linearity between pairs of descriptors (a value less than 0.5 indicated the absence of co-linearity between descriptors); and (3) the significance of the regression model (for a significance level of 5%). Internal validation of the MDF-SAR models was analyzed in the fourth step by using the Leave-one-out Analysis application [1].

The comparison between the MDF-SAR model and previously reported models was performed in the fifth step by using the Steiger's Z test at a significance level of 5% [18].

The prediction ability of the best performing MDF-SAR model was analyzed in the sixth step by using the Training vs. Test application [3]. There were analyzed twelve situations, starting with sample sizes in training set from twenty and increasing with one until thirty-one and corresponding sample sizes in test sets from twenty to nine.

RESULTS

Two MDF SAR models, one with two and one with four descriptors revealed to had estimation and prediction abilities. The MDF SAR models are:

- Model with two descriptors:

$$\hat{Y}_{2D}=1.74+1.01\cdot10^{-1}\cdot inPRlQg +3.10\cdot10^{-3}\cdot lPDMqMg \qquad (1)$$

where $\hat{Y}_{2D}$ is the estimated inhibition activity on CA I, and *inPRlQg*, *lPDMqMg* are the molecular descriptors used by the model, respectively.

- Model with four descriptors:

$$\hat{Y}_{4D}=1.14+8.79\cdot10^{-2}\cdot inPRlQg+2.43\cdot iAMRqQg +3.52\cdot10^{-3}\cdot lPDMoMg+1.04\cdot inMRkQt \qquad (2)$$

where $\hat{Y}_{4D}$ is the estimated inhibition activity on CA I, and *inPRlQg*, *lPDMoMg*, *iAMRqQg*, and *inMRkQt* are the molecular descriptors used by the model.

Table 4. Values of descriptors used in Eq(1) and Eq(2) and estimated activity by the models

Descr.	1	2	3	4	5	1,2	2,3,4,5
Abb.	lPDMqMg	inPRlQg	lPDMoMg	iAMRqQg	inMRkQt	$\hat{Y}_{2v}$	$\hat{Y}_{4v}$
cle_01	-202.70	$-7.35\cdot10^{-1}$	-195.88	$2.77\cdot10^{-1}$	$-2.44\cdot10^{-3}$	1.0402	1.0582
cle_02	-250.21	$-2.24\cdot10^{0}$	-238.16	$1.66\cdot10^{-1}$	$-6.29\cdot10^{-3}$	0.7410	0.5018
cle_03	-254.01	$-2.82\cdot10^{0}$	-242.16	$2.67\cdot10^{-1}$	$-2.78\cdot10^{-3}$	0.6708	0.6869
cle_04	-263.42	$-1.52\cdot10^{-1}$	-252.06	$2.03\cdot10^{-1}$	$-1.44\cdot10^{-2}$	0.9109	0.7179
cle_05	-274.68	$-9.50\cdot10^{-2}$	-261.01	$2.08\cdot10^{-1}$	$-2.14\cdot10^{-2}$	0.8818	0.6978
cle_06	-287.95	$-1.02\cdot10^{-2}$	-268.42	$1.70\cdot10^{-1}$	$-1.31\cdot10^{-2}$	0.8492	0.5934
cle_07	-261.07	$-2.62\cdot10^{-1}$	-249.10	$2.27\cdot10^{-1}$	$-1.21\cdot10^{-4}$	0.9071	0.7914
cle_08	-264.11	$-2.05\cdot10^{0}$	-251.74	$4.39\cdot10^{-1}$	$-3.06\cdot10^{-2}$	0.7178	1.1098
cle_09	-257.27	$-3.54\cdot10^{0}$	-245.86	$3.66\cdot10^{-1}$	$-7.27\cdot10^{-2}$	0.5879	0.7775
cle_10	-258.69	$-3.54\cdot10^{0}$	-247.21	$4.28\cdot10^{-1}$	$-2.54\cdot10^{-1}$	0.5840	0.7348
cle_11	-273.32	$-5.46\cdot10^{-1}$	-262.32	$2.01\cdot10^{-1}$	$-1.45\cdot10^{-3}$	0.8405	0.6573
cle_12	-278.37	$-1.79\cdot10^{0}$	-267.07	$2.78\cdot10^{-1}$	$-1.98\cdot10^{-2}$	0.6998	0.6995
cle_13	-287.56	$-1.62\cdot10^{0}$	-275.73	$2.98\cdot10^{-1}$	$-1.28\cdot10^{-2}$	0.6877	0.7380
cle_14	-231.46	$-8.07\cdot10^{-1}$	-222.47	$4.11\cdot10^{-1}$	$-9.85\cdot10^{-2}$	0.9439	1.1836
cle_15	-284.28	$-3.65\cdot10^{0}$	-272.15	$2.34\cdot10^{-1}$	$-1.23\cdot10^{-1}$	0.4934	0.3021
cle_16	-215.02	$-1.08\cdot10^{0}$	-200.56	$3.06\cdot10^{-1}$	$-2.03\cdot10^{-3}$	0.9673	1.0822
cle_17	259.68	$-3.12\cdot10^{-2}$	277.21	$4.05\cdot10^{-1}$	$-5.14\cdot10^{-1}$	2.5444	2.5639
cle_18	-302.78	$-4.64\cdot10^{-1}$	-290.90	$4.36\cdot10^{-1}$	$-3.14\cdot10^{-2}$	0.7575	1.1043
cle_19	-305.15	$-4.67\cdot10^{-1}$	-290.41	$3.05\cdot10^{-1}$	$-2.04\cdot10^{-3}$	0.7498	0.8172
cle_20	-274.02	$-5.43\cdot10^{-2}$	-261.23	$4.51\cdot10^{-1}$	$-2.52\cdot10^{-1}$	0.8879	1.0496
cli_01	-221.87	$-8.57\cdot10^{-1}$	-214.33	$3.55\cdot10^{-1}$	$-1.47\cdot10^{-2}$	0.9686	1.1587
cli_02	-260.01	$-9.99\cdot10^{-1}$	-246.66	$2.34\cdot10^{-1}$	$-8.95\cdot10^{-2}$	0.8360	0.6592
cli_03	-263.99	$-5.49\cdot10^{-2}$	-250.89	$2.96\cdot10^{-1}$	$-1.68\cdot10^{-1}$	0.9189	0.7985
cli_04	-273.93	$-6.34\cdot10^{-1}$	-261.36	$3.17\cdot10^{-1}$	$-1.74\cdot10^{-2}$	0.8297	0.9185
cli_05	-285.81	$-4.65\cdot10^{-2}$	-270.78	$2.67\cdot10^{-1}$	$-4.01\cdot10^{-2}$	0.8521	0.7904
cli_06	-299.80	$-3.03\cdot10^{-1}$	-278.50	$2.74\cdot10^{-1}$	$-7.58\cdot10^{-3}$	0.7829	0.7922
cli_07	-269.07	$-8.49\cdot10^{-1}$	-255.80	$3.45\cdot10^{-1}$	$-2.37\cdot10^{-1}$	0.8231	0.7576
cli_08	-267.76	$-6.52\cdot10^{-1}$	-253.99	$5.96\cdot10^{-1}$	$-1.21\cdot10^{0}$	0.8471	0.3804
cli_09	-267.44	$-9.06\cdot10^{0}$	-254.81	$2.22\cdot10^{-1}$	$-7.68\cdot10^{-4}$	0.0000	-0.0137
cli_10	-268.75	$-8.80\cdot10^{0}$	-256.08	$2.63\cdot10^{-1}$	$-3.81\cdot10^{-2}$	0.0224	0.0657
cli_11	-279.38	$-3.04\cdot10^{-1}$	-267.08	$3.93\cdot10^{-1}$	$-3.66\cdot10^{-1}$	0.8461	0.7488
cli_12	-284.65	$-5.57\cdot10^{-1}$	-272.07	$3.24\cdot10^{-1}$	$-8.62\cdot10^{-3}$	0.8043	0.9133
cli_13	-293.61	$-7.84\cdot10^{-1}$	-280.45	$2.67\cdot10^{-1}$	$-1.67\cdot10^{-2}$	0.7536	0.7156
cli_14	-248.94	$-2.03\cdot10^{-1}$	-239.27	$3.62\cdot10^{-1}$	$-3.21\cdot10^{-3}$	0.9506	1.1560
cli_15	-290.14	$-7.54\cdot10^{0}$	-276.65	$1.72\cdot10^{-1}$	$-1.86\cdot10^{-2}$	0.0834	-0.0968
cli_16	-204.03	$-2.66\cdot10^{-1}$	-187.76	$4.30\cdot10^{-1}$	$-1.13\cdot10^{-1}$	1.0835	1.3844
cli_17	323.41	$-7.08\cdot10^{-2}$	343.33	$5.16\cdot10^{-1}$	$-9.10\cdot10^{-1}$	2.7379	2.6513
cli_18	-301.52	$-8.11\cdot10^{-2}$	-288.13	$3.07\cdot10^{-1}$	$-4.16\cdot10^{-2}$	0.8000	0.8225
cli_19	-309.96	$-8.71\cdot10^{-1}$	-293.74	$3.06\cdot10^{-1}$	$-1.64\cdot10^{-2}$	0.6941	0.7578
cli_20	-279.37	$-6.77\cdot10^{-3}$	-265.09	$2.22\cdot10^{-1}$	$-1.81\cdot10^{-4}$	0.8761	0.7464

Table 5. Statistical parameter associated with the MDF SAR models

Parameter (abbreviation)	Value	
	MDF SAR model	
	2-D	4-D
Number of compounds (n)	40	40
Number of descriptors (v)	2	4
Correlation coefficient (r)	0.8975	0.9579
95% CI for correlation coefficient ($^{95\%}CI_r$)	[0.8133 - 0.9448]	[0.9212 - 0.9776]
Squared correlation coefficient (r^2)	0.8056	0.9175
Adjusted squared correlation coefficient (r^2_{adj})	0.7951	0.9081
Standard error of estimation (s_{est})	0.2426	0.1624
Fisher parameter (F_{est})	$77^{\dagger}$	$97^{\dagger}$
Cross-validation leave-one-out score ($r^2_{cv\text{-}loo}$)	0.7888	0.8911
Standard error of leave-one-out analysis (s_{loo})	0.2532	0.1869
Fisher parameter of loo analysis (F_{pred})	$69^{\dagger}$	$71^{\dagger}$
$r^2 - r^2_{cv\text{-}loo}$	0.0167	0.0264
r^2(*inPRlQg*, *lPDMqMg*)	0.0208	n.a.*
r^2(*inPRlQg*, *lPDMoMg*)	n.a.*	0.0216
r^2(*inPRlQg*, *iAMRqQg*)	n.a.*	0.0613
r^2(*inPRlQg*, *inMRkQt*)	n.a.*	0.0234
r^2(*lPDMoMg*, *iAMRqQg*)	n.a.*	0.1429
r^2(*lPDMoMg*, *inMRkQt*)	n.a.*	0.3123
r^2(*iAMRqQg*, *inMRkQt*)	n.a.*	0.4995

2-D: Two descriptors; 4-D: Four descriptors; † $p < 0.001$; *n.a. = not applicable

The calculated values of the descriptors and of the estimated inhibition activity on CA I obtained by the MDF SAR model with two ($\hat{Y}_{2v}$) and respectively with four descriptors ($\hat{Y}_{4v}$) are presented in Table 4.

Statistical parameters of the MDF SAR models from Eq(1) and Eq(2) are presented in Table 5 and 6.

The graphical representation of the inhibition activity on CA I of studied compounds estimated by Eq(2) versus measured is presented in Figure 1.

The correlated correlation analysis shown that the MDF SAR model with four descriptors obtained a correlation coefficient statistical significant greater comparing with the MDF SAR model with two descriptors (Steiger Z parameter = 3.28, significance of Steiger's parameter = $5.24\cdot10^{-4}$).

Internal validation of the MDF SAR model with four descriptors was analized through splitting the whole set into training and test sets using an original algorithm of randomization. The coefficients for each model, the number of compounds in training (n_{tr}) and test (n_{ts}) sets, the correlation coefficient obtained by training (r_{tr}) and by test (r_{ts}) sets with associated 95% confidence intervals (95% CI_{rtr} and 95% CI_{rts}), the Fisher parameter associated with training (F_{tr}) and test (F_{ts}) sets, and the Fisher's Z parameter of correlation coefficients comparison ($Z_{rtr-rts}$)) are presented in Table 7.

Table 6. Quality analysis of MDF SAR models

	SE	$r^2(Y, desc)$	t	$^{95\%}CI$	p-value
MDF SAR model with two descriptors					
Intercept	0.0845	n.a.*	20.62	[1.5715-1.9140]	$7.02\cdot10^{-22}$
inPRlQg	0.0174	0.2822	5.81	[0.0657-0.1360]	$1.14\cdot10^{-6}$
lPDMqMg	0.0003	0.6282	9.980	[0.0025-0.0037]	$4.84\cdot10^{-12}$
MDF SAR model with four descriptors					
Intercept	0.1295	n.a.*	8.799	[0.8768-1.4028]	$2.16\cdot10^{-10}$
inPRlQg	0.0119	0.2822	7.375	[0.0637-0.1121]	$1.26\cdot10^{-8}$
lPDMoMg	0.0002	0.6274	14.24	[0.0030-0.0040]	$3.95\cdot10^{-16}$
iAMRqQg	0.3812	0.2663	6.378	[1.6576-3.2055]	$2.46\cdot10^{-7}$
inMRkQt	0.1663	0.1299	6.249	[0.7013-1.3764]	$3.64\cdot10^{-7}$

SE = standard error; Y = log IC_{50}; desc = molecular descriptor; t = parameter of the Student test; p-value = t test significance; $^{95\%}CI$ = 95% confidence interval; * not applicable.

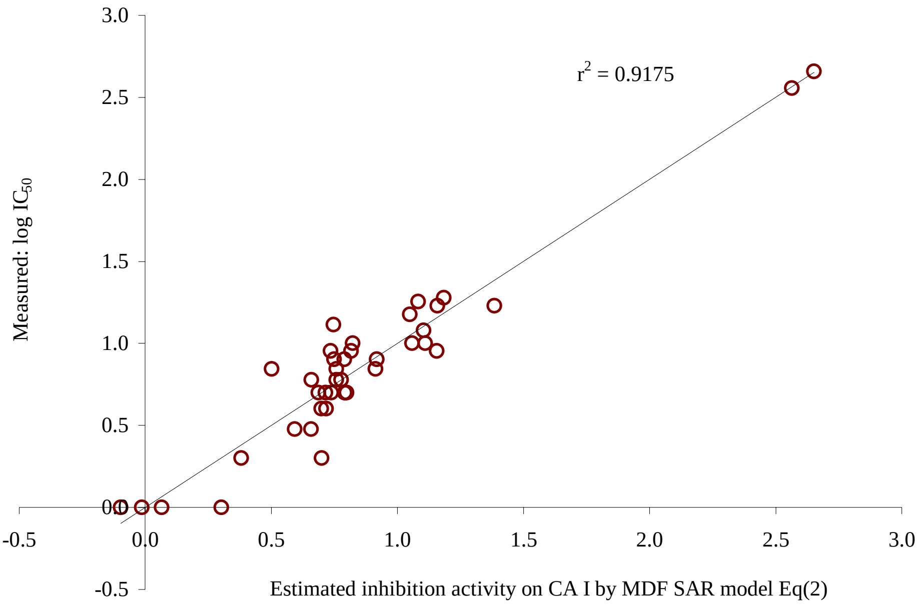


Figure 1. CA I inhibition activity on of studied compounds estimated by MDF SAR model with four descriptors

Table 7. Training versus test analysis: results

n_{tr}	20	21	22	23	24	25	26	27	28	29	30	31
a_0	1.257	1.078	0.899	1.069	0.689	1.455	0.909	1.193	1.169	1.088	1.094	1.205
a_1	$9.92\cdot10^{-2}$	$9.01\cdot10^{-2}$	$8.01\cdot10^{-2}$	$8.63\cdot10^{-2}$	$7.51\cdot10^{-2}$	$9.16\cdot10^{-2}$	$1.06\cdot10^{-1}$	$8.45\cdot10^{-2}$	$8.88\cdot10^{-2}$	$9.77\cdot10^{-2}$	$9.07\cdot10^{-2}$	$8.66\cdot10^{-2}$
a_2	$3.65\cdot10^{-3}$	$3.58\cdot10^{-3}$	$2.91\cdot10^{-3}$	$3.41\cdot10^{-3}$	$2.87\cdot10^{-3}$	$4.46\cdot10^{-3}$	$3.09\cdot10^{-3}$	$3.52\cdot10^{-3}$	$3.50\cdot10^{-3}$	$3.61\cdot10^{-3}$	$3.05\cdot10^{-3}$	$3.59\cdot10^{-3}$
a_3	2.209	2.72	2.527	2.456	3.317	2.095	2.841	2.213	2.332	2.694	2.167	2.221
a_4	1.154	1.152	0.606	0.826	1.256	0.45	0.849	0.984	1.015	1.163	0.977	1.005
r_{tr}	0.936	0.961	0.984	0.945	0.934	0.892	0.945	0.945	0.934	0.958	0.916	0.945
$^{95\%}CI_{rtr}$	[0.842, 0.975]	[0.905, 0.984]	[0.961, 0.993]	[0.873, 0.977]	[0.851, 0.971]	[0.766, 0.951]	[0.880, 0.975]	[0.882, 0.975]	[0.860, 0.969]	[0.911, 0.982]	[0.830, 0.960]	[0.887, 0.973]
F_{tr}	$27^‡$	$49^‡$	$129^‡$	$38^‡$	$33^‡$	$19^‡$	$44^‡$	$46^‡$	$39^‡$	$67^‡$	$33^‡$	$55^‡$
n_{ts}	20	19	18	17	16	15	14	13	12	11	10	9
r_{ts}	0.972	0.954	0.901	0.965	0.942	0.962	0.951	0.972	0.988	0.966	0.976	0.981
$^{95\%}CI_{rts}$	[0.929, 0.989]	[0.881, 0.982]	[0.750, 0.963]	[0.902, 0.987]	[0.837, 0.980]	[0.881, 0.988]	[0.848, 0.985]	[0.905, 0.992]	[0.957, 0.997]	[0.872, 0.991]	[0.897, 0.994]	[0.908, 0.996]
F_{ts}	$60^‡$	$34^‡$	$14^‡$	$38^‡$	$18^‡$	$6^†$	$18^‡$	$32^‡$	$70^‡$	$13^†$	$15^†$	$21^†$
$Z_{rtr-rts}$	1.23	0.27	$2.69^†$	0.65	0.18	1.53	0.15	0.88	$2.24^†$	0.28	1.49	1.18

a_0 = intercept; a_1 = inPRlQg; a_2= lPDMoMg; a_3 = iAMRqQg; a_4 = inMRkQt; ‡$p \leq 0.001$; †$0.001 < p < 0.05$

The results of comparison between previous reported models [19] and MDF SAR models are presented in Table 8.

Table 8. Results of comparison between previous reported models and MDF SAR models

QSAR - MDF SAR	Steiger Z	p-value
Model 1* - Eq(1)	0.582	0.2803
Model 2* - Eq(1)	1.041	0.1489
Model 1* - Eq(2)	2.563	0.0052
Model 2* - Eq(2)	2.965	0.0015

* Table 1,2

DISCUSSIONS

The inhibition activity on carbonic anhydrase I of substituted 1,3,4-thiadiazole- and 1,3,4-thiadiazoline-disulfonamides can be characterized starting from the complex information obtained from compounds structure. The sample of forty-studied compound was analyzed as a whole even if there was possibility to split it into two samples, as substituted thiadiazole disulfonamides and substituted thiadiazoline disulfonamides.

The MDF SAR model with two descriptors shown that the geometry (Eq(1), *inPRlQ**g*** and *lPDMqM**g***) of compounds is related with inhibition activity on CA I as well as partial charge (*inPRl**Q**g*), and atomic relative mass (*lPDMq**M**g*).

The goodness-of-fit of the MDF SAR model with two descriptors is sustained by the correlation coefficient and associated squared correlation coefficient (see Table 5). The cross-validation leave-one-out score ($r^2_{cv\text{-}loo}$) was higher than 0.7 (more, it had been decreasing by approximate 2% comparing with squared correlation coefficient), suggesting that according with Golbraigkh and Tropsha criteria the equation had predictive abilities [5]. Furthermore, the MDF SAR model with two descriptors is a stable model (the differences between squared correlation coefficient and cross-validation leave-one-out being equal with 0.0167, see Table 5). Almost eighty percent from variation of inhibition activity on CA I of substituted 1,3,4-thiadiazole- and 1,3,4-thiadiazoline-disulfonamides can be explained by the linear relationship with the variation of the two molecular descriptors used by the model (Eq(2)). The contribution of molecular descriptors to inhibition activity on CA I are equals with $1.01\cdot10^{-1}$ and $3.10\cdot10^{-3}$, respectively, these contributions being direct related with the activity of interest. These results suggest that the inhibition activity on CA I of studied compounds is of geometrical nature, being related with the partial charges and atomic relative mass of the compounds.

Analyzing the previous reported models (model 1 and model 2 from Table 1 and 2) and comparing them with the MDF SAR model with two descriptors, comparison applied on the correlation coefficients, it can be observed that there are not statistical differences between models (see Table 8). But, analyzing the number of variable used by previous reported models and MDF SAR model with two descriptors, it can be observed that the MDF SAR model obtained the same performances in estimation of inhibition activity on CA I with two descriptors comparing with previously reported model that used six (model 1), and five (model 2) variable, respectively.

Starting from the MDF SAR model with two descriptors and from its performances, the modeling process of inhibition activity on CA I of studied compound was run further and the multiple linear regression analysis identify a model with four descriptors. One descriptor from the MDF SAR model with two descriptors was finding again in the MDF SAR model with four descriptors.

The goodness-of-fit of four-varied model is sustained by the correlation coefficient that was of 0.9579 and its squared value that was of 0.9175. Almost ninety-two percent from the variation of inhibition activity on CA I of studied compounds can be explained by the linear relationship with the variation of the four molecular descriptors used by the model (Eq(2)). All contributions of molecular descriptors to inhibition activity had positive signs, marking out a direct relationship with the activity of interest. Looking at the name of descriptors it can be observed that the inhibition activity is on geometry (*inPRlQ**g***, *lPDMoM**g***, *iAMRqQ**g***) as well as topology (*inMRkQ**t***) nature, depend on atomic relative mass (*lPDMo**M**g*) and on the partial charges (*inPRl**Q**g*, *iAMRq**Q**g*, *inMRk**Q**t*) of the compounds.

The results of the cross-validation leave-one-out analysis sustain the predictive ability of the MDF SAR model with four descriptors. The difference between cross-validation leave-one-out score and the squared correlation coefficient was equal with 0.0264 (with almost 3% less comparing with squared correlation coefficient).

The power of the MDF SAR model with four descriptors in prediction of inhibition activity on CA I of studied compounds is sustained by the absence of multi-collinearity of descriptors used (see the squared correlation coefficients between pairs of descriptors, which always is less than 0.49, Table 5).

The comparison between MDF SAR models shown that the correlation coefficient obtained by the model with four descriptors is statistical significant greater comparing with the correlation coefficient obtained by model with two descriptors ($p < 0.001$). For this reason, the internal validation by splitting the sample in training and test sets was performed just for the MDF SAR model with four descriptors. As it can be seen from the results presented in Table 7, for all sample sizes in training and test sets the regression models were statistical significant and the correlation coefficients were not exceeded the 95% confidence interval of correlation coefficient obtained for the MDF SAR model with four descriptors. Furthermore, just in two cases out of twelve there were identified significant differences between correlation coefficients obtained in training and test sets: in one case the correlation coefficient obtained in test set was less than the one obtained in training set, while in the other case the correlation coefficient obtained in training set was less than one obtained in test set (see Table 7). The intercept of the regression models and the coefficients associated to molecular descriptors in regression equations in training versus test analysis (see Table 7), respected in the majority of the cases the 95% confidence intervals of the MDF SAR model with four descriptors (see Table 6 and 7).

Comparing the MDF SAR model with four variables (Eq(2)) with previous reported models (model 1 and 2 from Table 1 and 2) some remarks can be make. First remark refers the number of descriptors used in the models: previous reported models used in both cases more descriptors (six - model 1, respectively five - model 2) comparing with MDF SAR model with four descriptors. Second remark refers the squared correlation coefficients: both previous reported models had correlation coefficients statistical significant less comparing with MDF SAR model with four descriptors ($p < 0.006$, Table 8). Last but not lest, even if the cross-validation scores reported previously are greater than 0.5, the differences between squared correlation coefficient and cross-validation score, in both cases (model 1 and 2, Table 2), are greater than 10% (the predictive abilities could be in these conditions questionable). Note that, the values of the cross-validation leave-one-out scores are with 2% respectively 3% less than the values of squared correlation coefficients (see Table 5). Two out of four previously reported models (model 3 and 4, Table 1 and 2) where not considered for the analysis because these models used just twenty compounds as sample size, considering the thiadiazoles (model 2, Table 1 and 2) or thiadiazoline (model 4, table 1 and 2). More, the model 4 (Table 1) did not accomplish the Hawkins criterion of validation [11] ($n \geq 4\text{-}5 \cdot v$, where n is the sample size and v is the number of variables), the model taking into consideration seven variables (the sample size necessary in order to be a valid model must be 28, 35 respectively).

Further research are necessary in order to characterized the role of the MDF SAR model with four descriptors in development of new compound with inhibitory potencies on CA I from disulfonamides class. These investigations must be done on other disulfonamides than those included into the study. Based on the MDF SAR model with four descriptor and by the use of original software [2] the inhibition on CA I of new disulfonamides can be characterized and analyzed without any experiments. The steps necessary to be accomplished are: sketching out the molecular structure of compound by the use of HyperChem software, choosing from the list display by the software [2] the MDF SAR model, browsing the *hin file, predicting and displaying the inhibition activity on CA I of new compound.

Modeling the inhibition activity on carbonic anhydrase I of substituted 1,3,4-thiadiazole- and 1,3,4-thiadiazoline-disulfonamides by integration of complex structural information provide stable models with two and four descriptors allowing us to characterized the relationship between the compounds structure and inhibition activity on CA I.

The MDF SAR model with four descriptors shown that the inhibition activity on CA I of substituted 1,3,4-thiadiazole- and 1,3,4-thiadiazoline-disulfonamides is like to be of geometry and topology nature, being related with two atomic properties, represented by partial charge and relative atomic mass.

ACKNOWLEDGEMENT

Research was partly supported by UEFISCSU Romania through project ET36/2005.

REZUMAT

Relația între structura şi activitatea inhibitoare a anhidrazei carbonice tip I a fost studiată cu ajutorul familiei de descriptori moleculari MDF pe un set de patru zeci de substituenți ai 1,3,4-tiadiazol- şi 1,3,4-tiadiazolin- disulfonamide. Familia de descriptori moleculari a fost generată pornind de la informația obținută din structura compuşilor şi astfel s-au calculat descriptorii. Ecuațiile structură-activitate s-au obținut folosind setul de descriptori calculați. Modelele reprezentative, cu cea mai bună performanță în estimarea activității biologice, au fost identificate folosind coeficientul de determinare, parametrul Fisher şi probabilitatea asociată acestuia. A fost analizată abilitatea de prezicere a două modele multivariate; de asemenea au fost comparați coeficienții de corelație ai modelelor obținute cu coeficienții de corelație obținuți de alte modele raportate în literatura de specialitate. Rezultatele arată că MDF SAR este un instrument util în caracterizarea activității inhibitoare a anhidrazei carbonice de tip I de către compuşii de substituție ai 1,3,4-tiadiazol- şi 1,3,4-tiadiazolin- disulfonamidelor.

REFERENCES

1. ***, 2005© - Leave-one-out Analysis, Virtual Library of Free Software [cited 2006 December]. Available from: http://vl.academicdirect.org/molecular_topology/mdf_findings/loo/
2. ***, 2005© - MDF SAR Predictor, Virtual Library of Free Software [cited 2006 December]. Available from: http://vl.academicdirect.org/molecular_topology/mdf_findings/sar
3. ***, 2005© - Training vs. Test Experiment, Virtual Library of Free Software [cited 2006 December]. Available from:

http://vl.academicdirect.org/molecular_topology/mdf_findings/qsar_qspr_s/
4. AGRAWAL VK, BANO S, SUPURAN CT, KHADIKAR PV, 2004 - QSAR study on carbonic anhydrase inhibitors: aromatic/heterocyclic sulfonamides containing 8-quinoline-sulfonyl moieties, with topical activity as antiglaucoma agents, Eur J Med Chem, 39(7), p. 593-600.
5. BASAK SC, NIEMI GJ, REGAL RR, VEITH GD, 1987 - Topological Indices: their Nature, Mutual Relatedness, and Applications, Mathematical Modelling, 8, p. 300-305.
6. de LEVAL X, ILIES M, CASINI A, DOGNE JM, SCOZZAFAVA A, MASINI E, MINCIONE F, STARNOTTI M, SUPURAN CT, 2004 - Carbonic anhydrase inhibitors: synthesis and topical intraocular pressure lowering effects of fluorine-containing inhibitors devoid of enhanced reactivity, J Med Chem, 47(11), p. 2796-804.
7. DEUTSCH HF, 1987 - Carbonic anhydrases, Int. J. Biochem, 19, p. 101-113.
8. DIUDEA M, GUTMAN I, JÄNTSCHI L, 2002 - Molecular Topology, 2nd Edition, Nova Science, Huntington, New York, USA.
9. GUPTA SP, KUMARAN S, 2005 - A quantitative structure-activity relationship study on some aromatic/heterocyclic sulfonamides and their charged derivatives acting as carbonic anhydrase inhibitors, J Enzyme Inhib Med Chem, 20(3), p. 251-259.
10. HANSCH C, 1969 - A Quantitative Approach to Biochemical Structure-Activity Relationships, Acct Chem Res, 2, p. 232-239.
11. HAWKINS DM, 2004 - The Problem of Overfitting, J Chem Inf Comput Sci, 44, p. 1-12.
12. HEWETT-EMMETT D, TASHIAN RE, 1996 - Functional diversity, conservation, and convergence in the evolution of the alpha-, beta-, and gamma-carbonic anhydrase gene families, Mol Phylogen Evol, 5, p. 50-77.
13. JÄNTSCHI L, 2005 - Molecular Descriptors Family on Structure Activity Relationships 1. The review of Methodology, Leonardo Electronic Journal of Practices and Technologies, 4(6), p. 76-98.
14. JÄNTSCHI L, BOLBOACA SD, 2006 - Modelling the Inhibitory Activity on Carbonic Anhydrase IV of Substituted Thiadiazole- and Thiadiazoline- Disulfonamides: Integration of Structure Information, Electronic Journal of Biomedicine, Red UniNet, 4(2), p. 22-33.
15. JÄNTSCHI L., BOLBOACA S, (2007) - Results from the Use of Molecular Descriptors Family on Structure Property/Activity Relationships, Int J Mol Sci, 8(3), p. 189-203.
16. JÄNTSCHI L, KATONA G, DIUDEA MV, 2000 - Modeling Molecular Properties by Cluj Indices, Commun Math Comput Chem (MATCH), 41, p. 151-188.
17. JÄNTSCHI L, UNGURESAN ML, BOLBOACA SD, 2005 - Complex Structural Information Integration: Inhibitor Activity on Carbonic Anhydrase II of Substituted Disulfonamides, Applied Medical Informatics, 17, p. 12-21.
18. STEIGER JH, 1980 - Tests for comparing elements of a correlation matrix, Psychol Bull, 87, p. 245-251.
19. SUPURAN CT, CLARE BW, 2001 & 2002 - Orbital symmetry in QSAR: some Schiff's base inhibitors of carbonic anhydrase. SAR QSAR Environ Res., 12, p. 17-29 (2001), Erratum in: SAR QSAR Environ Res,13, p. 755 (2002).
20. SUPURAN CT, CLARE BW, 1999 - Carbonic anhydrase inhibitors - Part 57: Qunatum chemical QSAR of a group of 1,3,4-thiadiazole- and 1,3,4-thiadiazoline disulfonamide with carbonic anhydrase inhibitory properties, Eur J Med Chem, 34, p. 41-50.
21. TRIPP BC, Smith K, Ferry JG, 2001 - Carbonic Anhydrase: New Insights for an Ancient Enzyme, J Biol Chem, 276(52), p. 48615-48618.

AUTHOR'S ADDRESS

Sorana Daniela BOLBOACĂ: PhD, MSc Inf, MD
"Iuliu Hatieganu" University of Medicine and Pharmacy, 6 Louis Pasteur Str., 400349 Cluj-Napoca, Cluj, Romania,
http://sorana.academicdirect.ro,
sbolboaca@umfcluj.ro

Lorentz JÄNTSCHI: PhD, Chem, MSc Agr,
Technical University of Cluj-Napoca, Romania, 103-105 Muncii Bvd., 400641 Cluj-Napoca, Cluj, Romania, http://lori.academicdirect.org,
lori@academicdirect.org